APEC-ATCWG Workshop (2008)
Capacity Building for Risk Management Systems on Genetic Resources
71~78

Establishing a Taiwan Biodiversity Information Network and Its Integration with Germplasm Databanks

Kwang-Tsao Shao[1], Sheng-Chung Huang[2], Shu Chen[2], Yung-Chang Lin[3], Kun-Chi Lai[3], Burke Chih-Jen Ko[3], Lee-Sea Chen[3], Alan Jieluen Yang[3]

[1.3.] Biodiversity Research Center, Academia Sinica, Taipei
[2.] Taiwan Agricultural Research Institute, Taichung

Abstract

With the establishment of the “Biodiversity Promotion Plan” by the Executive Yuan, the National Science Council commissioned Academia Sinica to integrate Taiwan’s biodiversity databases. Subsequently Academia Sinica created TaiBNET in 2002 and TaiBIF in 2003. The Agricultural Research Institute’s National Plant Genetic Resources Center, which is in charge of integrating data on agriculture, forestry, fishery, animal husbandry, and germplasm, began collaborating with the Catalog of Life of TaiBNET in 2004. Through a linker, i.e., the scientific name, data in these two databases can be matched against each other. Because the TaiBIF is cooperating with GBIF, all germplasm data in Taiwan can now be accessed globally.

Introduction

Prior to 2001, biodiversity databases were established and located at different governmental and academic institutions and NGOs with no integration among them. At most, some links to certain related sites were provided. But then users had to search for the information they needed from the beginning home page. Integration work finally began in 2001, especially after the launch of the 1st five-year National Digital Archives Program (NDAP) which includes the natural heritage of specimen and species information. The “Biodiversity Promotion Plan” was also approved and executed by the Executive Yuan in the same year. That year, Taiwan joined the Global Biodiversity Information Facility (GBIF) as an Associate

[1] Executive Director/Research Fellow, Biodiversity Research Center, Academia Sinica.
E-mail: Zoskt@gate.sinica.edu.tw

Participant so we were able to use the same metadata of the Darwin Core and DiGIR as the exchange platform to link local databases to global databases. From 2002 to 2004, The National Science Council (NSC) commissioned us to establish the Catalog of Life website (TaiBNET) with a team of more than 100 taxonomists in Taiwan. The national portal of the GBIF in Taiwan (TaiBIF) was established in 2003 under sponsorship of the NSC and the Council of Agriculture (COA). In 2007, we merged the TaiBIF into the 2nd phase of the NDAP, now called TELDAP, because international collaboration is one of the foci for this national program.

The Biodiversity Action Plan has two missions. (1) Mission D1201 – Establish biodiversity data exchange mechanisms as well as establish and integrate related databases of government agencies and institutions; and periodically add content. The executive agency is the NSC and co-execution agencies include the COA, Academia Sinica, Ministry of the Interior, Ministry of Economic Affairs, Council of Indigenous Peoples, Department of Health, Ministry of Education, Environmental Protection Administration, and Ministry of Transportation and Communications. (2) Mission D1301 – Establish national germplasm banks; and conduct research, conservation, and utilization of agriculture, forestry, fishery, animal husbandry, wildlife, and microorganism genetic resources. The COA is in charge, and co-sponsors are the NSC, Academia Sinica, Ministry of Education, and Ministry of Economic Affairs. This is also why Academia Sinica and the National Plant Genetic Resources Center (NPGRC) need to work together.

Using scientific names to integrate all biodiversity information

A correct and up-to-date biological species checklist is the most important step for integrating biodiversity databases. This is because each species has only one valid name, and the name is the only keyword or linker in all biological databases. Through the scientific name, specimen information can be accessed and queried, as can genetic data (Barcode of Life, BOL), Tree of Life (TOL), ecological distributional data, and updated species information in the Encyclopedia of Life (EOL). For work on the Catalog of Life (COL) in Taiwan, we just held a “2008 Workshop: Research and Status of Species Diversity in Taiwan” in August of this year. In November, we will publish both a hardcopy and a CD-ROM of the 1st official national species checklist. As to the BOL, Academia Sinica is a member of the Consortium of Barcode of Life organizations. In September 2007, we successfully hosted the 2nd International Barcode of Life Conference in Taipei. So far, about 300,000 barcodes for 30,000 species have

been put on the internet for public access.

As to wild animals, the Forestry Bureau of the COA is promoting a team project for collecting wildlife tissue samples for cryobanking and barcoding. Academia Sinica is in charge of establishing this database and has already accumulated about 3000 species and 8000 lots of tissue samples preserved in liquid nitrogen along with their voucher specimens. In this database, a user can learn how many species and which species have been collected by various institutions, including Academia Sinica (fishes), the Taiwan Endemic Species Research Institute, Taiwan Normal University (birds and mammals), Taipei City Zoo, NTNU (amphibians and reptiles), National Museum of Natural Sciences (invertebrates and insects), and Taiwan Livestock Research Institute (central bank).

The first 30,000 pages of species information, one species a page, of the online EOL have been made available to the public. The goal is to achieve a total of 1.8 millions pages by 2017. We would like to collaborate with the EOL in providing pages of Taiwanese endemic species from our own “Taiwan EOL”.

The TaiBIF is the national portal for the GBIF

Figure 1 shows the homepages of the Chinese and English TaiBIF website. The Life Science Department of the NSC and COA are sponsors of updating the species checklist and integrating *Digital fauna and flora*, as well as collecting data on ecological distributions; whereas NDAP funding helps integrate all contents collected from NDAP sub-projects and cross-agency integration. This website can also search for endemic species images, species information, and geographical distribution range, and link to other biodiversity databases in Taiwan, including regulations, institutions, NGOs, national parks, etc. In total, 11,775 species have been compiled in the Fauna and Flora of Taiwan in the TaiBIF. An additional 2000~3000 species are expected to be added by the end of the year.

For biodiversity-related literature, theses, symposia, research projects, and reports in Taiwan, the Science and Technology Policy Research and Information Center, National Applied Research Laboratories has collected about 31,000 records. Biodiversity-related news and activities in Taiwan are updated on a weekly basis. The website can be linked to the website of the Environmental Information Center. The geographical distribution map for each species and the specimens collected can be displayed on a GIS map by dots or 10 x 10-km grids. These data can also be exported to Google maps. Since the TaiBIF is one of the GBIF data providers, when the latitude and longitude data of these specimens are submitted to

GBIF, the locations can then be accessed on the GBIF website or other global databases such as FishBase or OBIS.

Fig. 1. TaiBIF website in Chinese and English.

TaiBNET is the catalog of life (COL) in Taiwan

To collaborate with the COL of Species 2000, we mainly use the TaiBNET website. This database generally is for integrating and updating Taiwanese native and introduced species checklists. So far, we have accumulated seven kingdoms, 51 phyla, 2614 families, and about 48,000 native species, not including introduced species. The organisms can be searched for by the classification system of kingdom phylum, class, order, and family. The statistics on the current species occurring in Taiwan in each family, order, class, and phylum are provided and are continually being updated. When a user clicks on a family, all species belonging to that particular family are listed. The species name can also be searched for using taxon names, key words, or strings. When a particular species that a user is interested in is selected, detailed information on the species is given. There are also links to other databases in Taiwan

or global databases so that further information on that particular species can be accessed.

Currently, TaiBNET has collected about 35,967 synonyms, 24,313 Chinese common names, 3417 taxonomic names, and 2953 corresponding Chinese names at the rank of family and above. Collecting synonyms is important because scientific names are often changed. Only a very few taxonomic experts know which name is valid or invalid, and users often cannot find relevant data when using old names. Since TaiBNET features synonyms, it automatically finds the correct name if a user keys in an invalid one. All data in TaiBNET are open to the public, so users, if they wish, can download data using either xml or csv format.

Many species or strains deposited at the agriculture, forestry, fisheries, and animal livestock institutes, as well as the Bioresources Collection and Research Center (BCRC) are introduced species or strains, not native species. As a result, we decided to compile lists of introduced, cultural, and naturalized species in the TaiBNET as well. Otherwise, it would not be possible to link to and retrieve the information for many non-native species. To encourage users to provide their own data, we have also prepared an interface webpage so that users can upload photos, species information, and synonyms.

Another database that the TaiBNET hosts is the experts name list, mainly for experts in taxonomy and ecology. Through this database, users can find the expert of one special taxon by choosing the expert's name, specialty, organization, or occupation as the condition of the cross search. The resulting page contains both Chinese and English information. Another advantage of having a good database is the capability to verify the correctness of the data. The TaiBNET now offers users the feature of double-checking whether the species composition or list contains spelling or typing errors, or whether the species is valid or not. By pasting the name list on the webpage of the TaiBNET, a user will receive a list of possible errors for correction. This function can significantly improve data quality.

Integration of the TaiBNET/TaiBIF with germplasm databanks in Taiwan

Figure 2 shows how we have integrated different databases from different agencies. Basically, we use the species name or the GIS distribution (GPS data) as the primary key or linker to directly access raw data or webpages of other databases. The TaiBNET and TaiBIF are the two portals for this cross-agency integration in Taiwan.

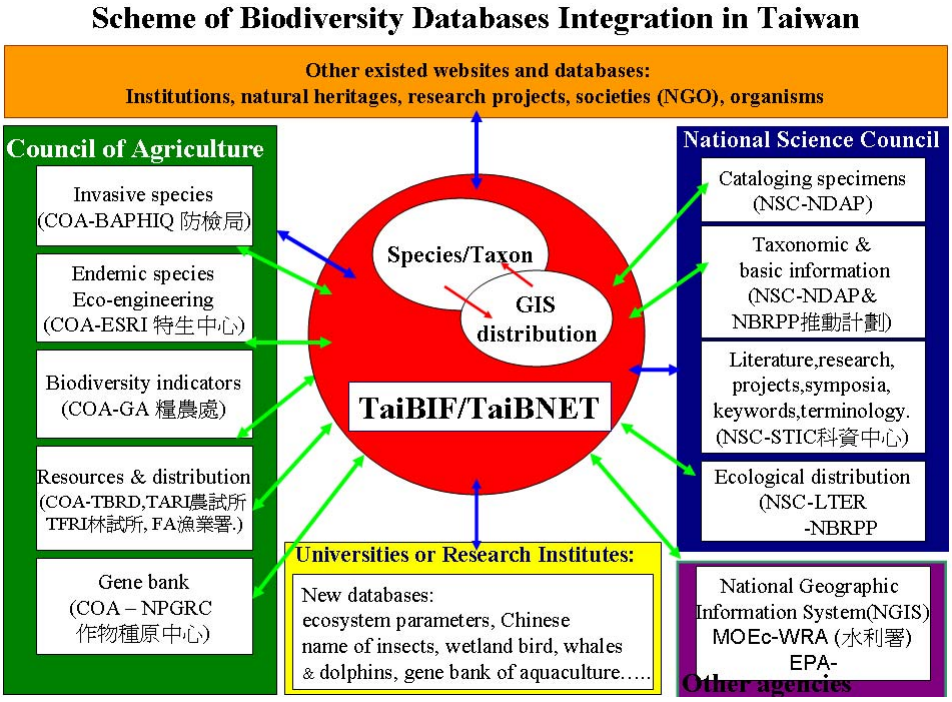

Fig. 2. Integration of databases from various agencies.

The NPGRC is in charge of coordinating germplasm databanks at various institutes of the COA. So far, the Agricultural Research Institute has collected about 1434 species which belong to 189 families and 781 genera. The total numbers of data down to strains or lines include 68,108 passport data, 27,581 records of characterization data, and 4002 images. In order to facilitate the exchange of plant germplasm information among scientists and breeders, the COA sponsored a 3-year project to integrate all germplasm data from four institutes and integrate those data with the TaiBIF and TaiBNET during 2006~2008. The first step was to compare the species lists between germplasm banks and the TaiBNET and then modify the classification system and match the species names. Table 1 shows the number of species matched up between each institute and the TaiBNET/TaiBIF, including the BCRC. Table 2 is a list of institutions and databases that have been integrated with the TaiBNET so that all information can be directly linked or mutually exchanged with each other. The BCRC, the largest and central bank for microorganisms in Taiwan, provides data on 104 species. So far, we only have a list of germplasm data from the Forestry Research Center, because they have not provided URLs of their 108 species. The Fishery Research Institute has gamete characteristics on 40 freshwater aquaculture fishes.

Table 1. No. of genera/ No. of species matching up between various institutes and the TaiBNET/TaiBIF

TaiBNET & TaiBIF include agriculture, forestry, fishery and animal husbandry information

	Genus	Species
Bioresources Collection and Research Center	53	104
Agricultural Research Institute(NPGRC)	428	675
Taiwan Forest Research Institute	77	108
Livestock Research Institute	16	21
Livestock Research Institute – Poaceae	32	52
Livestock Research Institute – Fabaceae	26	38

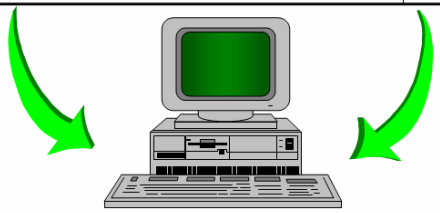

Table 2. A list of institutions and databases that have integrated with the TaiBNET

675 species at National Plant Genetic Resources Center are linked to TaiBNET through scientific names

TaiBNET 台灣物種名錄 Catalogue of Life in Taiwan

首頁 物種名錄 專家名錄 相關資料 訊息服務

本站推薦連結的網站(以學名直接與該站連結)

網站中文名稱	網站英文名稱	連結筆數
U.S. National Parasite Collections	U.S. National Parasite Collections	83
Manter Laboratory Of Parasitology	Manter Laboratory Of Parasitology	2
台灣大型甲殼類資料庫	The Large Crustacean of Taiwan	319
台灣本土植物影像資料庫	Images Database of Native Plants in Taiwan	223
台灣貝類資料庫	The Taiwan Malacofauna Database	3421
台灣魚類資料庫	The Fish Database of Taiwan	2917
台灣博物館海藻資訊網	Seaweeds Net of Taiwan	156
生物資源保存及研究中心	Bioresources Collection and Research Center (BCRC)	104
自然科學博物館自然與人文網站	The Digital Museum of Nature and Culture (National Museum of Natural Science)	1733
國立海洋生物博物館生物典藏與數位資訊網	National Museum of Marine Biology Aquarium Digital Network	596
世界魚庫 FishBase	FishBase	2917
台灣昆蟲學會	Taiwan Entomological Society	1
楊懿如的青蛙學堂	Frog School	32
台灣大蚊網站	Tipulidae in Taiwan	12
Discover Life	Discover Life	23488
農試所作物種原中心	National Plant Genetic Resources Center	675
國立台灣大學台灣植物誌電子書	Flora of Taiwan 2nd eBooks (National Taiwan University)	6
台灣原住民族生物學誌	台灣原住民族生物學誌	613

In addition to the TaiBNET, all of the germplasm data mentioned above can be searched for on the TaiBIF website as well. Due to the fact that the TaiBIF is collaborating with the GBIF, all germplasm data in Taiwan can now be accessed globally through the GBIF or any other national portal.

Future work and perspectives

Work remaining to be accomplished includes: establishing a common name list in English; increasing the number of Chinese common names; adding and linking more species names in germplasm banks; linking more webpages with information on species and lines (strains); setting up links to global germplasm data banks; and for TaiBNET, increasing the number of introduced and cultivated species included in the Catalog of Life in Taiwan.